PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 22 July 15, 1936 Number 7

ON THE PERIODIC ANALYTIC CONTINUATIONS OF THE CIRCULAR ORBITS IN THE RESTRICTED PROBLEM OF THREE BODIES

By Aurel Wintner

Department of Mathematics, The Johns Hopkins University

Communicated June 11, 1936

The following considerations, suggested by Part I, Section 9, of Birkhoff's recent work on the restricted problem of three bodies,[1] deal with the stability character of the periodic orbits which represent isoenergetic analytic continuations of Kepler's circular solutions ($\mu = 0$) for positive values of the mass μ of the perturbing body. Let the units be chosen in the usual way. Then, if a denotes the radius of the orbit of the perturbed body in the limiting case $\mu = 0$, Hill's ratio of periods is $m = 1/(n-1)$, where n is the positive or negative square root of a^{-3} according as the path is direct or retrograde in the siderial coördinate system. One has to exclude not only the values $m = 0$ and $m = -1$, which belong to $a = 0$ and $a = \infty$, but also the value $m = -1/2$, which belongs to the orbit ($a = 1$) of the perturbing body. Small positive values of m represent the case of the lunar theory, while non-vanishing integral values of m belong to the critical commensurabilities of minor planets.

It is known[2] that if m is neither an integer nor the number $m = -1/2$ excluded above, then there is a sufficiently small positive $\lambda = \lambda(m)$ such that there exists for $0 < \mu < \lambda(m)$ a series of isoenergetic periodic solutions which depend on the parameter μ in a regular analytic manner and go over into the Keplerian circle belonging to the given value of m, as $\mu \to 0$. Let $C = C(m)$ be the Jacobian energy constant of this family. Let i denote any fixed closed m-interval containing neither an integral value of m nor the value $m = -1/2$. Then there exists a positive number $\lambda = \lambda_i$ such that $\lambda(m)$ may be chosen greater than λ_i for every m in i. This, and even more, has been proved by Birkhoff[3] in the case where i lies between $m = 0$ and $m = 1$, i.e., between the range of lunar orbits and the commensurability of Hecuba. If i has any other position, the proof for the existence of a positive λ_i requires but obvious modifications. Thus one obtains for every fixed positive $\mu \leqq \lambda_i$ a one-parametric family of periodic

orbits, the varying parameter of the family being Hill's ratio m or, what amounts to the same thing, the Jacobian energy constant C. The family is obtained when, μ having a fixed value, m describes the m-interval i, and so C a corresponding C-region which will be denoted by j. Let $F_j(\mu)$ denote the family of periodic motions thus obtained for a fixed $\mu \leqq \lambda_i$.

Now it is possible to choose $\lambda_i > 0$ so small that if μ has any fixed value not greater than λ_i, the orbits which constitute $F_j(\mu)$ depend on their parameter C in a regular analytic manner, when C varies on j. This may be shown by an easy analysis of the proof[2] which assures the existence of a positive $\lambda = \lambda(m)$ for a fixed m.

The variational equation of the normal displacements, i.e., the equation of Jacobi for the periodic solution which belongs to the given values of μ and C, is easily obtained.[4] If $\mu = 0$, this equation is[1] $d^2z/dt^2 + n^2z = 0$, if the motion is referred to the synodical coördinate system. It follows[4] that for sufficiently small μ the equation of Jacobi is of the form

$$d^2z/dt^2 + \{n^2 + f(\mu, C, t)\}z = 0, \tag{1}$$

where $f(\mu, C, t)$ is continuous in μ, C, t together, and that $f(\mu, C, t) \to 0$ as $\mu \to 0$. The function $f(\mu, C, t)$ has with respect to t the same period $T = T(\mu, C)$ as the periodic solution of the family $F_j(\mu)$ to which (1) belongs. Furthermore, $f(\mu, C, t) = f(\mu, C, -t)$, since this periodic solution is known[5] to be symmetric with respect to the axis of syzygies. Thus

$$f(\mu, C, t) = \sum_{k=0}^{\infty} a_k \cos(2\pi kt/T), \tag{2}$$

where $a_k = a_k(\mu, C) \to 0$ as $\mu \to 0$, since $f \to 0$. Also

$$T = T(\mu, C) \to 2\pi m \text{ as } \mu \to 0. \tag{3}$$

This is clear from the definition of Hill's ratio m.

A direct calculation shows that if C is fixed and $\mu > 0$, then the small coefficient $a_1 = a_1(\mu, C)$ is of a lower order in μ than $a_2, a_3, \ldots$ are together, and that, in particular, a_1 is distinct from zero for small positive values of μ. Thus it is seen from (2) and (3) that, when μ is very small, the solutions of

$$d^2z/dt^2 + \{[n^2 + a_0] + a_1 \cos(2\pi t/T)\}z = 0 \tag{4}$$

give a first approximation to those of (1) in view of the continuity theorems of linear differential equations. On placing $u = 2\pi t/T$, it is seen from (3), where $m = 1/(n-1)$, that (4) goes over into

$$d^2z/du^2 + \{[(m+1)^2 + b_0] + b_1 \cos u\}z = 0, \tag{5}$$

where $b_0 \to 0$ and $b_1 \to 0$ as $\mu \to 0$. Furthermore, b_1 does not vanish for small positive μ, since the same holds for a_1. It is understood that b_0 and

b_1 depend on both μ and C, and that (5) gives for small μ an approximation to (1) which is uniform for all C in any fixed C-region j.

Since $b_0 \to 0$ and $0 \neq b_1 \to 0$ as $\mu \to 0$, one may write (5) in the form

$$d^2z/dv^2 + (q^2 + p \cos 2v)z = 0, \qquad (6)$$

where $v = {}^1/_2u$, while q and p are two non-vanishing real numbers. Let h be a non-vanishing integer. Consider for a moment Mathieu's equation (6) with given parameters q and p. If q varies on a small interval S containing $q = h$ in its interior, while p is fixed and has a sufficiently small absolute value distinct from zero, then S will be the sum of three intervals S_1, S_2, S_3 such that S_1 and S_3 are separated by S_2 and the characteristic exponent of (6) is of the elliptic (stable) or hyperbolic (unstable) type according as q is or is not in S_2, it being understood that the stability character is indifferent at the end-points of S_2, the multiplier being $+1$ or -1 according as h is even or odd. This phenomenon of small intervals S_2 of formal instability in case of Mathieu's equation was discovered by Poincaré,[6] who gave the proof for the case where the integer h contained by S is small. Poincaré suspected, and Strutt[7] recently proved, that the result holds for any h.

These results on (6) are applicable to (5), since $b_0 = b_0(\mu, C)$ and $b_1 = b_1(\mu, C) \neq 0$ tend, as $\mu \to 0$, to zero uniformly for all C in any fixed C-region j. The values of m which correspond to integral values h of q are halves of integers, since

$$(2m + 2)^2 + 4b_0 \to (2m + 2)^2$$

in view of (5) and (6), where $v = {}^1/_2u$. Actually, one must confine himself to the halves of odd integers which are, in addition, distinct from $-{}^1/_2$, since the m-interval i is supposed to be free of the points representing integral values of m or $m = -{}^1/_2$. Since (4) is identical with (5), the intervals of formally unstable character exist, of course, for (4) also.

Now let j be a C-region of the type defined above and let j be chosen so long that the corresponding m-region i contains the half of an odd integer in its interior. Then, if μ has a sufficiently small fixed positive value, j is the sum of three intervals j_1, j_2, j_3 such that j_1 and j_3 are separated by j_2 and that the characteristic exponent of periodic orbit which belongs to the energy constant C in the family $F_j(\mu)$ is of the instable or the stable type according as C does or does not lie on the segment j_2 of j, it being understood that the end-points of j_2 belong to periodic solutions of indifferent stability character, the multiplier being -1. The existence of the interval j_2 of formally hyperbolic character has been obtained above not with regard to the exact equation (1) of Jacobi but with regard to the approximation (4) to (1). However, a_0 and a_1 are of lower order in μ than all neglected coefficients a_2, aj, ... of (2) together, and this holds uniformly

for all C on j. Hence it is shown by careful, though straightforward, estimates of the errors introduced by neglecting $a_2, a_3, \ldots$ that, if the fixed positive number μ is sufficiently small, the decomposition of j into one hyperbolic and two elliptic regions of the characteristic exponent cannot be lost by the transitions from (4) to the exact Jacobian equation (1).

Accordingly, the situation is as follows. Consider a C-region j such that the end-points of the corresponding m-region are very near to two subsequent integers and contain neither an integer nor $m = -1/2$. Then, if μ has a sufficiently small fixed positive value and the parameter C of the family $F_j(\mu)$ describes the range j, the periodic orbit is first stable, then unstable, finally again stable, stability and instability being meant in the sense of the characteristic exponents. If, for instance, the pair of subsequent integers is (0, 1), so that j is the family joining the lunar orbits with orbits of the Hecuba type, the region in which the non-vanishing characteristic exponent is of the hyperbolic type will lie in the neighborhood of the commensurability $m = 1/2$ of Hestia.

As pointed out above, the family $F_j(\mu)$, where μ is fixed, depends on its parameter C in a regular analytic manner. Consequently, there is an exchange of formal stability at the end-points of the middle part j_2 of the range $j = j_1 + j_2 + j_3$, although the end-points of j_2 are not branch points of the family. Conversely, there exist families having a branch point at which an exchange of formal stability does not take place.[8] Correspondingly, Poincaré[9] did not prove his well-known rule as a general theorem which does not allow any exception, but rather as a useful principle which is valid for certain generic, but not for all possible, cases.

If $\mu = 0$, the multiplier is ± 1 if and only if m is the half of an integer, even or odd, while the multiplier is complex and of absolute value 1, i.e., of the stable type, for the remaining values of m. This is the reason[10] that for small fixed positive μ the characteristic exponents will be of the stable type except in the neighborhood of values C which belong for $\mu = 0$ to integral values of $2m$. If m is an integer distinct from -1 and 0, then[11] $\lambda(m) = 0$; actually,[3] the case $m = 0$ (or, rather, the case $a \to 0$) is as harmless as possible. The case where m is the half of an integer but not an integer and distinct from $-1/2$ has been treated above with the result that in some small region in the neighborhood of any of these commensurabilities one must have formal instability, hence instability proper, if $\mu > 0$ is small. The theorem may be illustrated by the numerical results of Darwin.[12]

[1] G. D. Birkhoff, *Pisa Annali*, **4**, 267–306 (1935); **5**, 9–50 (1936).

[2] Cf. T. Levi-Civita, *Annali di Mat.* (3) **5**, 286–289 (1901); G. D. Birkhoff, *Rend. Palermo*, **39**, chap. 11 (1915).

[3] G. D. Birhoff, loc. cit.,[2] chap. 12.

[4] Cf. A. Wintner, *Sächs. Sitzber.*, **82**, 345–354 (1930).

[5] H. Poincaré, *Méth. Nouv.*, **1** (1892), chap. III; G. D. Birkhoff, loc. cit.[1]

[6] H. Poincaré, op. cit., **2** (1893), chap. XVII. The figure given by Poincaré assumes that p is positive. However, the case of a negative p may be reduced to the case of a positive p, as seen from (6) by writing $v + 1/2\pi$ for v.

[7] M. J. O. Strutt, *Math. Ann.*, **99**, 625–628 (1928).

[8] Cf. A. Wintner, *Bull. Astr.*, **9**, 251–253 (1936).

[9] H. Poincaré, op. cit., **3**, 346–351 (1899).

[10] H. Poincaré, op. cit., **3**, 343–344 (1899).

[11] Cf. E. Hölder, *Sächs. Sitzber.*, **83**, 179–184 (1931).

[12] G. D. Darwin, *Scientific Papers*, **4** (1911), Part I. Cf. H. Poincaré, op. cit., **3**, 352–361 (1899).

ON THE STRUCTURE OF NATIVE, DENATURED, AND COAGULATED PROTEINS

BY A. E. MIRSKY* AND LINUS PAULING

GATES CHEMICAL LABORATORY, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIFORNIA

Communicated June 1, 1936

In this paper a structural theory of protein denaturation and coagulation is presented. Since denaturation is a fundamental property of a large group of proteins, a theory of denaturation is essentially a general theory of the structure of native and denatured proteins. In its present form our theory is definite and detailed in some respects and vague in others; refinement in regard to the latter could be achieved on the basis of the results of experiments which the theory suggests. The theory (some features of which have been proposed by other investigators) provides a simple structural interpretation not only of the phenomena connected with denaturation and coagulation which are usually discussed (specificity, solubility, etc.) but also of others, such as the availability of groups, the entropy of denaturation, the effect of ultra-violet light, the heat of activation and its dependence on pH, coagulation through dehydration, etc.

I. The experimental basis upon which the present theory rests will be briefly described.

1. The most significant change that occurs in denaturation is the loss of certain highly specific properties by the native protein. Specific differences between members of a series of related native proteins and specific enzymatic activities of native proteins disappear on denaturation, as the following observations demonstrate:

(*a*) Many native proteins can be crystallized and the crystal form is characteristic of each protein. No denatured protein has been crystallized.

(*b*) Most native proteins manifest in their immunological properties a high degree of specificity, which is diminished by denaturation.[1]

(*c*) The native hemoglobins of closely related animal species can be distinguished from each other by differences in crystal form, solubility,[2] gas affinities, positions of absorption bands, and other properties.[3] On the other hand, in the denatured hemoglobins some of these properties, the positions of the absorption bands, for example, can be subjected to precise measurement, and it is found that differences between the various hemoglobins have disappeared.[4]

(*d*) A number of enzymes have recently been isolated as crystalline proteins. When these proteins are denatured their enzymatic activity vanishes. In pepsin and trypsin, where an especially careful study has been made of this phenomenon, there is a close correlation between loss of activity and formation of denatured protein.[5]

2. Striking changes in the physical properties of a protein take place during denaturation. At its isoelectric point a denatured protein is insoluble, although the corresponding native protein may be quite soluble. It was the loss of solubility that first drew attention to the phenomenon of denaturation, and denaturation is now usually defined by the change in solubility. The denatured protein after precipitation has taken place is called a coagulated protein, the process of coagulation being considered to include both denaturation and aggregation of denatured protein in the form of a coagulum.[6] If the denatured protein is dissolved, by acid, alkali, or urea, the solution is found to be far more viscous than a solution of native protein of the same concentration.[7]

3. Changes in the availability of sulfhydryl, disulfide, and phenol groups appear as a consequence of denaturation. All of the SH and S-S groups found in a protein after hydrolysis can be detected in a denatured protein even before hydrolysis, while in the corresponding native protein only a fraction of these groups is detectable. In native egg albumin no SH or phenol groups are detectable. In other native proteins (hemoglobin, myosin, proteins of the crystalline lens, for example) some groups can be detected; new groups appear when the protein is made more alkaline, although not alkaline enough to cause denaturation, and then disappear when the original pH is restored. In all the different ways of coagulating a protein a close correlation between appearance of groups and loss of solubility has until recently been observed.[8] Now, however, it has been found that when myosin is rendered insoluble by drying (or when water is removed by freezing) there is no change in availability of its SH groups.[9] Furthermore, if the insoluble myosin is treated with a typical denaturing agent, such as heat or acid, all of the SH groups in the protein become available, although no change in solubility is observed. Hitherto protein coagulation due to dehydration has not been distinguished from coagulation

caused by other agents, but the tests for availability show that in coagulation by dehydration the change in protein constitution is distinctly different from that caused by any of the known protein coagulating agents. In our theory of coagulation both types of change will be considered. Probably both types occur biologically. It has been suggested that when light converts visual purple, a conjugated protein, into visual yellow the former is denatured in much the same way as it is by heat, alcohol, or acid.[10] And it has been shown that in the course of fertilization and in the rigor of muscle a coagulation of protein similar to that caused by dehydration takes place.[11,12]

4. In a list of the large number of different agents, with apparently little in common, that cause denaturation are heat, acid, alkali, alcohol, urea, salicylate, surface action, ultra-violet light, high pressure. The temperature coefficient of heat denaturation of many proteins (egg albumin, ferrihemoglobin, trypsin, etc.) is about 600 for a rise in temperature of ten degrees, and from this the energy of activation can be calculated. On either side of a point between the isoelectric point and neutrality the temperature coefficient of denaturation by heat is diminished.[13] A dry preparation of egg albumin is not readily denatured by heat.[6]

5. The denaturation of certain proteins, notably hemoglobin, serum albumin, and trypsin, is reversible.[14] From the effect of temperature on the equilibrium constant, the heat of reaction and the entropy change can be calculated. In the denaturation of hemoglobin by salicylate and in the denaturation of trypsin by heat or acid, the equilibrium between native and denatured protein is not affected by changes in the total concentration of protein present,[15] from which it can be inferred that in the denaturation of hemoglobin and trypsin by these agents no change in molecular weight takes place. Since typical denaturation can occur without change in the molecular weight, it is important to distinguish between denaturation and the changes in particle size observed by Svedberg. Although denaturation can occur without change in molecular weight, under certain conditions, as in the denaturation of myosin by urea, denaturation may be accompanied by depolymerization. Weber found the molecular weight of myosin to be of the order of a million and that of myosin in urea to be thirty-five thousand.[16] On the other hand, it is unlikely that depolymerization is always accompanied by denaturation, for under some of the conditions of depolymerization described by Svedberg it is improbable that denaturation (loss of specificity, loss of solubility, or appearance of previously inaccessible groups) takes place. In the case of hemocyanin, for example, decomposition into products $^1/_2$ and $^1/_{16}$ of the size of the original molecule readily occurs, and these smaller particles seem to have the properties of native hemocyanin.[17]

6. The shape of the native protein molecule appears to have little sig-

nificance for an understanding of denaturation. Denaturation occurs in the spherical molecules of egg albumin and hemoglobin, in the elongated particles of soluble myosin[18] and (as indicated by SH groups becoming detectable) even in myosin that has been formed into insoluble fibres by drying.[9]

7. In certain conjugated proteins stability of the protein and presence of the prosthetic group are related. In hemoglobin, the yellow oxidizing ferment of Warburg, and visual purple, it is necessary to denature the protein in order to detach the prosthetic group with the use of present methods.[10,19] After removal of the prosthetic group it is possible to reverse the denaturation of globin and the protein of the oxidizing ferment, but these native proteins are more unstable (with respect to denaturation) than they are when conjugated with their prosthetic groups. In hemoglobin the ease of denaturation depends upon the state of the prosthetic group. Carbon monoxide hemoglobin, for example, is less readily denatured by heat, acid, or alkali than are oxyhemoglobin and ferrihemoglobin.[20] In visual purple presence of the prosthetic group causes the protein to be denatured by visible light. Denaturation in this case appears to be reversible.[10]

II. Our conception of a native protein molecule (showing specific properties) is the following. The molecule consists of one polypeptide chain which continues without interruption throughout the molecule (or, in certain cases, of two or more such chains); this chain is folded into a uniquely defined configuration, in which it is held by hydrogen bonds[21] between the peptide nitrogen and oxygen atoms and also between the free amino and carboxyl groups of the diamino and dicarboxyl amino acid residues.

We shall not enter into a long discussion of the precise configurations of native proteins, about which, indeed, little reliable information is available. From the x-ray investigations of Astbury[22] and his collaborators it seems probable that in most native proteins the polypeptide chain, with the extended or one of the contracted configurations discussed by Astbury, folds back on itself in such a way as to form a layer in which peptide nitrogen and oxygen atoms of adjacent chains are held together by hydrogen bonds; several of these layers are then superposed to form the complete molecule, the bonds between layers (aside from the continuation of the polypeptide chain from one layer to the next) being hydrogen bonds between side-chain amino and carboxyl groups. In general not all of the side chain groups will be used in forming bonds within the molecule; some will be free on the surface of the molecule.

The importance of the hydrogen bond in protein structure can hardly be overemphasized. No complete review of the large amount of recent work on this bond is available; we shall mention only the most striking of its properties.[21] The hydrogen bond consists of a hydrogen atom which bonds two electronegative atoms together (F, O, N), the hydrogen atom lying between the two bonded atoms. The bond is essentially electrostatic in nature. The bonded atoms are held more closely together than non-bonded atoms, the N-H-O distance being about 2.8 Å. The energy of a strong hydrogen bond is 5000 to 8000 cal. per mole, the lower value being approximately correct for an N-H-O bond as

in proteins. Side-chain bonds in proteins we consider to involve usually an amino and a carboxyl group, the nitrogen atom forming a hydrogen bond with each of two oxygen atoms and holding also one unshared hydrogen atom. In acid solutions hydrogen bonds may be formed between two carboxyl groups, as in the double molecules of formic acid.[28]

The characteristic specific properties of native proteins we attribute to their uniquely defined configurations.

The denatured protein molecule we consider to be characterized by the absence of a uniquely defined configuration. As the result of increase in temperature or of attack by reagents (as discussed below) the side-chain hydrogen bonds are broken, leaving the molecule free to assume any one of a very large number of configurations. It is evident that with loss of the uniquely defined configuration there would be loss of the specific properties of the native protein; it would not be possible to grow crystals from molecules of varying shapes, for example, nor to distinguish between closely related proteins when the molecules of each protein show a variability in configuration large compared with the differences in configuration of the different proteins.

Strong support of this view of the phenomenon of denaturation is provided by the known difference in entropy of native and denatured proteins, which is about 100 E. U. for trypsin, the entropy of the denatured form being the greater.[5] This very large entropy difference cannot be ascribed to a difference in the translational, vibrational, or rotational motion, but must be due to a difference in the number of accessible configurations. It corresponds to about 10^{20} accessible configurations for a denatured protein molecule. The large entropy of denaturation thus shows clearly that the phenomenon of denaturation consists in the change of the molecules of the native protein to a much less completely specified state.

The magnitude of the heat change, entropy change, and activation energy of denaturation (about 30,000 cal./mole, 100 E. U., and 150,000 cal./mole, respectively, for trypsin) can be interpreted in terms of our hydrogen-bond picture of the protein molecule. We consider the native protein molecule to be held in its definite configuration by side-chain hydrogen bonds, about fifty in number for this protein (corresponding to about twenty-five amino and twenty-five carboxyl side chains), each with a bond energy of about 5000 cal./mole, as in simpler systems. The activation energy of 150,000 cal./mole shows that in order for the molecule to lose its native configuration about thirty of the bonds must be broken. Some of the side-chain groups then again form hydrogen bonds; the heat of denaturation shows that on the average there are about six fewer such bonds in the denatured molecule than in the native molecule. (The activation energy for transition of a denatured protein molecule from one configuration to another is without doubt smaller than the activation energy of denaturation, as it

involves breaking a smaller number of hydrogen bonds, so that at a temperature at which the rate of denaturation is appreciable the denatured molecule would run rapidly through its various configurations.) The magnitude of the entropy of denaturation also fits into our picture; it corresponds to the number of configurations obtained by forming hydrogen bonds at random between about twenty amino side chains and twenty carboxyl side chains.

The reagents which cause denaturation are all substances which affect hydrogen-bond formation. Alcohol, urea, and salicylate are well-known hydrogen-bond-forming substances; they form hydrogen bonds with the protein side chains, which are thus prevented from combining with each other and holding the protein in its native configuration. Acids act by supplying protons individually to the electronegative atoms which would otherwise share protons, and bases by removing from the molecule the protons needed for hydrogen-bond formation. This conception provides an explanation of the facts that the isoelectric point of a protein shifts toward the neutral point on denaturation[24] and that the pH at which the activation energy for denaturation has its maximum value is in general not at the isoelectric point of the native protein, but between this point and the neutral point.[25] In the native protein molecule of the usual type some amino and carboxyl side-chain groups are paired together by forming hydrogen bonds. The acid-base properties of the molecule are in the main determined by the groups which are left free. On denaturation some of the paired groups are freed, amino and carboxyl in equal numbers, and in consequence the isoelectric point of the denatured protein is shifted toward neutrality. We have pictured the process of denaturation as involving the rupture of a large number of hydrogen bonds, to form a labile activated molecule. This labile molecule is stabilized by action of base on its free carboxyl groups or of acid on its free amino groups, the activation energy thus having a maximum at a pH value between the neutral point and the isoelectric point of the native protein.

The action of ultra-violet light must be different. It is not possible to formulate a reasonable mechanism whereby a quantum of light can break twenty or thirty hydrogen bonds. Instead the light must attack the molecule in a different place, probably breaking the main polypeptide chain after absorption in a tyrosine or other phenolic residue, as suggested by Mitchell.[26] That this occurs is indicated by the observation[27] that after illumination with ultra-violet light in the cold denaturation occurs only on warming, though then at a lower temperature than without illumination; it is clear that illumination in the cold causes a break in the molecule, which, however, is restrained to configurations near to its native configuration by the side-chain hydrogen bonds. (That some loosening of the molecule occurs is shown by the observation of an increase in the available

groups in egg albumin after illumination in the cold.) On warming, these bonds are broken; because of the break in the molecule, however, it can be denatured "in parts," and hence at a lower temperature than before illumination. We predict that it will be found that denaturation by illumination with ultra-violet light is in general not reversible.

In a conjugated protein the prosthetic group plays a part in holding the molecule in the native configuration (hemoglobin, yellow oxidizing ferment, visual purple). It is possible in such a protein for reversible denaturation to take place after absorption of light by the prosthetic group, no permanent damage being done to the molecule.

In a protein coagulum side-chain hydrogen bonds hold adjacent molecules together. Native proteins do not coagulate because most of the side chains are in protected positions inside of the molecule; denatured proteins (at the isolectric point) do coagulate because they have a larger number of free side chains and because in the course of time, as the molecule assumes various configurations, all of the side chains become free. The increase of viscosity of protein solutions on denaturation we attribute to the change from the compact configuration of the native protein molecules to more extended configurations.

A native protein molecule of small molecular weight may have free side chains so arranged as to permit it to combine with similar molecules to form a polymer with properties differing little from those of the small molecules. The observations of Svedberg and his collaborators indicate that this is the case for hemocyanin, casein, and certain other proteins.

Our theory of denaturation leads to definite predictions regarding the availability of groups to attack by reagents. In the large compact molecule of a native protein, of the order of magnitude of 50 Å in diameter and having the same structure as every other molecule of the protein, all groups would be protected from attack by reagents except those on the surface or near the surface. After denaturation of the protein, however, the molecule (in solution) would in the course of time assume various configurations, and every group in the molecule would become available to attack. These statements are in complete agreement with the experimental results regarding sulfhydryl, disulfide, and phenol groups mentioned above. The observation that the number of available groups is increased when the solution is made alkaline, though not alkaline enough to cause denaturation, shows that under these conditions some of the hydrogen bonds in the protein molecule are broken, causing it to assume a configuration somewhat more open than its original configuration.

Although a denatured protein molecule in a coagulum is not free to assume all configurations, being restrained by bonds to its neighbors, in general its configuration will be so open as to make all groups accessible to attack. However, as mentioned above, there is no change in the avail-

ability of groups in myosin when it is rendered insoluble by drying or when water is removed by freezing. We interpret this as showing that *coagulation of this type is not accompanied by denaturation; that is, the molecules do not lose their uniquely defined configurations.* On dehydration of the native protein the surface side chains of adjacent molecules form bonds sufficiently strong to produce a coagulum of native protein molecules. It would be expected that the mechanical bolstering effect of adjacent molecules in this coagulum would aid the molecules to retain their native configurations; this is observed to be the case for egg albumin and some other proteins, which in this state are denatured by heat only at a temperature considerably higher than before dehydration.[6]

Some features of the theory of protein structure discussed above have been suggested before. Many investigators have correlated the specific properties of native proteins with definitely specified molecular configurations and the loss of specific properties on denaturation with change in configuration. Astbury and his collaborators in particular, in discussing their x-ray investigations, which have provided so much valuable information on protein structure, have stated[22] that dehydration and temperature denaturation lead to destruction of the original special configuration of the native protein, giving a débris consisting simply of peptide chains. Our picture agrees with theirs except in regard to the effect of dehydration, which we believe to consist primarily in the coagulation of molecules with essentially unchanged structure. The hydrogen bonds which we postulate to exist between side-chain carboxyl and amino groups (each nitrogen atom attached by hydrogen to two oxygen atoms, with the distances N-H-O equal to 2.8 Å) are a refinement of the side-chain salt-like linkages discussed by Astbury. Our picture of the spreading of protein films on water, involving the unfolding of compact molecules to layers one amino-acid-residue thick, is not essentially different from that of Gorter and Neurath, who have suggested unfolding in connection with film formation, though not with denaturation in general.[28]

In this paper we have discussed in some detail a very drastic change in configuration of protein molecules, that connected with denaturation. We have pointed out that the large entropy of denaturation provides strong support for our suggested structures of native and denatured protein molecules, and that many other phenomena can also be interpreted in a simple way from this point of view. There also has been put forth recently some evidence that small changes in configuration of proteins occur, which play an important part in protein behavior. Thus Northrop and his collaborators have prepared an enzymatically inactive protein which on slight hydrolysis by trypsin is transformed into native trypsin,[29] and another which is similarly transformed by pepsin into native pepsin,[30] and Gorter[31] has shown that myosin spreads on water only after slight hydrolysis. The

structural interpretation of these phenomena can be made only after further experimental information is available.

* On leave of absence from the Hospital of the Rockefeller Institute.

[1] Zinsser, H., and Ostenberg, Z., *Proc. New York Path. Soc.*, **14**, 78 (1914).

[2] Landsteiner, K., and Heidelberger, M., *Jour. Gen. Physiol.*, **6**, 131 (1933).

[3] Barcroft, J., *The Respiratory Function of the Blood*, Part II, Cambridge Univ. Press, 1928.

[4] Anson, M. L., and Mirsky, A. E., *Jour. Physiol.*, **60**, 50 (1925).

[5] Northrop, J. H., *Jour. Gen. Physiol.*, **13**, 756 (1930); **16**, 33, 323 (1932).

[6] Chick, H., and Martin, C. J., *Jour. Physiol.*, **40**, 404 (1910); **43**, 1 (1911–12); **45**, 61, 261 (1912–13).

[7] Anson, M. L., and Mirsky, A. E., *Jour. Gen. Physiol.*, **15**, 341 (1932).

[8] Heffter, A., *Mediz. nat. Arch.*, **1**, 81 (1907); Arnold, V., *Zeit. Physiol. Chem.*, **70**, 300 (1911); Mirsky, A. E., and Anson, M. L., *Jour. Gen. Physiol.*, **18**, 307 (1934–35); **19**, 427, 439 (1936).

[9] Mirsky, A. E., unpublished experiments.

[10] Wald, G., *Jour. Gen. Physiol.*, **19**, 351 (1935); Mirsky, A. E., *Proc. Nat. Acad. Sci.*, **22**, 147 (1936).

[11] Mirsky, A. E., *Jour. Gen. Physiol.*, **19** (1936), in press.

[12] Mirsky, A. E., unpublished experiments.

[13] Lewis, P. S., *Biochem. Jour.*, **20**, 978, 984 (1927); Loughlin, W. J., *Biochem. Jour.* **27**, 1779 (1933).

[14] Anson, M. L., and Mirsky, A. E., *Jour. Gen. Physiol.*, **9**, 169 (1925); *Jour. Phys. Chem.*, **35**, 185 (1931).

[15] Anson, M. L., and Mirsky, A. E., *Jour. Gen. Physiol.*, **17**, 393, 399 (1934).

[16] Weber, H. H., *Ergeb. d. Physiol.*, **36**, 109 (1934).

[17] Svedberg, T., *Science*, **79**, 327 (1934).

[18] von Muralt, A., and Edsall, J. T., *Jour. Biol. Chem.*, **89**, 315 (1930).

[19] Theorell, H., *Biochem. Zeitschr.*, **278**, 263 (1935).

[20] Hartridge, H., *Jour. Physiol.*, **44**, 34 (1912).

[21] Latimer, W. M., and Rodebush, W. H., *Jour. Am. Chem. Soc.*, **42**, 1419 (1920); Pauling, L., *Proc. Nat. Acad. Sci.*, **14**, 359 (1928); Sidgwick, N. V., *The Electronic Theory of Valency*, Oxford University Press, 1929; Bernal, J. D., and Megaw, H. D., *Proc. Roy. Soc.*, **A151**, 384 (1935).

[22] Astbury, W. T., and Street, A., *Phil. Trans. Roy. Soc.*, **A230**, 75 (1931); Astbury and Woods, H. J., *Ibid.*, **A232**, 333 (1933); Astbury and Sisson, W. A., *Proc. Roy. Soc.*, **A150**, 533 (1935); Astbury and Lomax, R., *Jour. Chem. Soc.*, **1935**, 846; Astbury, Dickinson, S., and Bailey, K., *Biochem. Jour.* **29**, 2351 (1935); see also Meyer, K. H., and Mark, H., *Der Aufbau der Hochpolymeren Organischen Naturstoffe*, Akademische Verlagsgesellschaft, M. B. H., Leipzig, 1930.

[23] Pauling, L., and Brockway, L. O., *Proc. Nat. Acad. Sci.*, **20**, 336 (1934).

[24] Heidelberger, M., and Pederson, K. O., *Jour. Gen. Physiol.*, **19**, 95 (1935); Pederson, K. O., *Nature*, **128**, 150 (1931).

[25] Loughlin, W. J., *Biochem. Jour.*, **27**, 1779 (1933).

[26] Mitchell, J. S., *Nature*, **137**, 509 (1936).

[27] Bovie, W. T., *Science*, **37**, 373 (1913).

[28] Gorter, E., and Grendel, F., *Proc. Acad. Sci. Amsterdam*, **29**, 371 (1926); Neurath, H., *J. Phys. Chem.*, **40**, 361 (1936).

[29] Kunitz, M., and Northrop, J. H., *Science*, **80**, 505 (1934).

[30] Herriott, R. M., and Northrop, J. H., *Ibid.*, **83**, 469 (1936).

[31] Gorter, E., *Nature*, **137**, 502 (1936).

INVERSIONS IN THE THIRD CHROMOSOME OF WILD RACES OF DROSOPHILA PSEUDOÖBSCURA, AND THEIR USE IN THE STUDY OF THE HISTORY OF THE SPECIES

BY A. H. STURTEVANT AND TH. DOBZHANSKY

W. G. KERCKHOFF LABORATORIES OF THE BIOLOGICAL SCIENCES, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated June 8, 1936

Genetic studies showed several years ago that the third chromosomes of wild strains of Drosophila pseudoöbscura often carry suppressors of crossing-over. The salivary gland chromosome technique has made it possible to demonstrate not only that these are inverted sections, but also that there are many different inversions present in wild populations inhabiting different geographical regions. So far we have found at least fourteen different gene-sequences in wild stocks, and have found that in most geographical regions several sequences are present, though no single sequence appears to occur throughout the range of the species. There are a number of problems raised by these facts, most of which need further study; the present account is to be regarded only as a preliminary note.

One sequence has been arbitrarily designated as "standard," since it is the one in which most mutant genes are known and is also the only sequence that has been found in both of the cross-sterile races, *A* and *B*. This sequence is commonest in race *A* on the Pacific coast, from British Columbia to Lower California, but is found sporadically also in western Montana, western Nebraska, central Texas, central New Mexico (Carizozo) and southern Arizona (Santa Rita Mts.). In race *B* the standard sequence is predominant in the Sierra Nevada and the Coast Range south of the San Francisco Bay, California, but it has been found also in a single strain from the Olympic peninsula, Washington. Several other sequences represent single inversions with respect to the standard one; the "Arrowhead" sequence, found throughout the range of race *A* in British Columbia and the United States; the "Pikes Peak" sequence, found in race *A* from western Nebraska south to central Texas and southeastern Arizona; the "Klamath" sequence, found in race *B* from British Columbia south to Mono county, California; and the "Sequoia" sequence, known only in race *B* from Sequoia National Forest in California.

All the other known sequences differ from the standard by more than a single inversion; one of them, "Cowichan" in race *B* from Vancouver Island, differs from Klamath by a single inversion; "Chiricahua II," in race *A* from southeastern Arizona, differs from Arrowhead by a single inversion. All the others differ from all those so far named by more than a single inversion.

These facts give some information as to the historical relationships of the sequences concerned. If the sequence of the loci in the standard be arbitrarily designated *ABCDEFGH*, we may represent Klamath as *ADCBEFGH*, Arrowhead as *ABEDCFGH*, and Pikes Peak as *ABCGFEDH*. It is clear that the standard sequence could give rise to any of the three others by a single inversion, or that any of the three could give rise to standard by a single inversion. But it is equally clear that Arrowhead cannot have given rise directly (i.e., by a single inversion) either to Pikes Peak or to Klamath. If Arrowhead was the original, ancestral sequence then Pikes Peak and Klamath are related to it (and to each other) only through the standard, unless one assumes the occurrence of a complex

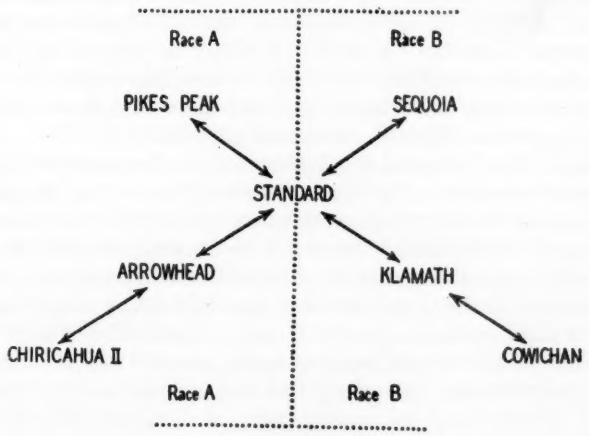

FIGURE 1

series of coincidences (repeated breakages of the chromosome in exactly the same places). This method enables us to construct a phylogeny of the types concerned, though it does not give us a method of determining the direction in which that phylogeny should be read. For the types named the relations are shown in figure 1.

There is no direct method of determining which is the original sequence here; but if any one be selected, then the whole course of the evolution is thereby fixed. However, it seems likely that the standard sequence is, in fact, the ancestral one, for the following reasons among others: (1) it is centrally located in the diagram, being the only type to which several others are directly related; (2) it is the only sequence found in both races (*A* and *B*); (3) it has a wide geographical range, being equalled in this respect only by the Arrowhead sequence.

None of the other chromosomes of Drosophila pseudoöbscura shows a

diversity of sequence at all comparable to that of the third chromosome. Tan[1] has shown that races *A* and *B* differ by having three inversions with respect to each other—one in each arm of the *X*-chromosome, and one in the second chromosome. We have shown[2] that the right limb of the *X* often has an inversion associated with the "sex-ratio" gene, but the total number of sequences known for this limb is three. In the course of the present study three additional inversions have been found in wild strains: one in the second chromosome of a few strains from western New Mexico, one in the second chromosome of an *A* strain from Monterey county, California, and one in a fourth chromosome from the Mexican plateau. There are thus known from wild strains the following numbers of sequences: the left limb of *X*, two; the right limb of *X*, three; the second chromosome, four; the third, fourteen; the fourth, two. The reason for these relations remains unknown; it should, however, be pointed out that the third chromosome which shows maximum variation within the species pseudoöbscura is also the chromosome which shows maximum differences from the closely related species, Drosophila miranda.[3]

As stated, many localities yield several third chromosome sequences from a single collection. This is perhaps related to the fact that from 15 to 20 per cent of the third chromosomes tested from wild stocks have been found to carry lethal genes. This raises the question as to whether such stocks are in a condition approaching that of balanced lethals. The problem is now being studied; but two facts make us hesitate to conclude that diversity of sequence and frequency of lethals are causally related. (1) In northern and western New Mexico and in the parts of Colorado, Utah and Arizona near the point where these four states come together, extensive tests have failed to show any sequence other than Arrowhead; yet lethals were recovered here in a proportion not noticeably less than that found elsewhere. The number of tests for lethals was, however, small. (2) Preliminary tests indicate that the frequency of lethals is about the same in the second chromosome as in the third, despite the much greater uniformity in sequence in the former (complete uniformity in race *A* stocks from the regions studied).

[1] Tan, C. C., "Salivary Gland Chromosomes in the Two Races of Drosophila pseudoöbscura," *Genetics*, **20** (1935).

[2] Sturtevant, A. H., and Dobzhansky, Th., "Geographical Distribution and Cytology of 'Sex-Ratio' in Drosophila pseudoöbscura and Related Species," *Genetics*, **21** (1936).

[3] Dobzhansky, Th., and Tan, C. C., "Studies on Hybrid Sterility. III. A Comparison of the Gene Arrangement in Two Species, Drosophila pseudoöbscura and Drosophila miranda." In press.

THE EFFECT OF TEMPERATURE ON THE FREQUENCY OF SOMATIC CROSSING-OVER IN DROSOPHILA MELANOGASTER

BY CURT STERN AND VIOLET RENTSCHLER

DEPARTMENT OF ZOÖLOGY, UNIVERSITY OF ROCHESTER

Communicated June 15, 1936

The frequency of crossing-over between chromosomes during gametogenesis is known to be subject to temperature influences. It seemed desirable to investigate a possible effect of temperature on *somatic* crossing-over.

Methods.—The occurrence of somatic crossovers in *Drosophila melanogaster* heterozygous for suitable factors results in the appearance of mosaic spots.[1] The number of spots is a measure of the frequency of somatic crossing-over. Females with the zygotic constitution Minute-*n*/yellow white singed[3] were inspected. They appear normal except for the presence of the short Minute-*n* bristles. Somatic crossing-over to the right of *Mn* leads to cells of the constitution *y w sn*^{3}/*y w sn*^{3} which appear as patches of yellow singed[3] setae or of white eye-facets according to the body region. Other types of crossovers result in different spots. Out of a total of 120 spots found during the course of this study 118 were derived from crossing-over to the right of *M-n*, 1 from crossing-over between *y* and *sn*^{3} and 1 from a double crossover on both sides of *sn*^{3}. The distribution of spots over different individuals was as follows: 109 flies with 1 spot each, 4 with 2 spots, 1 with 3 spots. The search for spots was restricted to an inspection of the macrochaetae and the eye only. All experiments were carried out by the junior author.

The flies were raised at three different temperatures—17° ± 0.5°C., 25° ± 0.1°C. and 29.5° ± 0.1°C. 2–3 *Mn*/+ females and 3–5 *y w sn*^{3} males were mated together in each culture bottle.

Results.—In order to compare the frequency of somatic crossing-over at 25° and 17° four sets of cultures were raised. They were started at different times from parental stocks which had not been highly inbred. Accordingly it is not surprising that the percentages of mosaic spots varied considerably (table 1). However, there was a high degree of correlation between the frequencies at the two temperatures for the different sets. In no case is the difference in percentage of crossing-over significant. This is borne out by the total of these experiments. At 25° a frequency of 12.7 ± 1.7 per cent was observed as compared to 11.4 ± 1.6 per cent at 17°. The difference between the two values is smaller than its standard error.

A significant difference was found in comparing somatic crossing-over at 25° and 30°. In both sets of experiments of this group a rather low percentage of mosaic spots at 25° was encountered, namely, 7.2 and 8.5 per

cent. The values at 30° were still lower, namely, 2.0 and 2.2 per cent. The total frequencies were 7.8 ± 1.6 at 25° and 2.1 ± 0.9 at 30°. The difference between these values is about 3.2 times its standard error.

A corroboration of this result is seen in an unpublished experiment performed in our laboratory by Mr. William James. His work was done in an

TABLE 1

NUMBER OF MOSAIC SPOTS IN MINUTE-*n* FLIES REARED AT DIFFERENT TEMPERATURES

SET	T°	NUMBER OF CULTURES	TOTAL M-N FLIES	TOTAL MOSAICS	% MOSAICS
1	25°	3	132	19	14.4
	17°	4	84	9	10.7
2	25°	3	129	11	8.5
	17°	4	135	11	8.1
3	25°	2	49	6	12.2
	17°	3	73	9	12.3
4	25°	5	85	14	16.5
	17°	5	84	14	16.7
Total 1–4	*25°*	*13*	*395*	*50*	*12.7 ± 1.7*
	17°	*16*	*376*	*43*	*11.4 ± 1.6*
5	25°	6	139	10	7.2
	30°	6	150	3	2.0
6	25°	3	142	12	8.5
	30°	3	92	2	2.2
Total 5–6	*25°*	*9*	*281*	*22*	*7.8 ± 1.6*
	30°	*9*	*242*	*5*	*2.1 ± 0.9*
W. J.	25.5°	19	345	16	4.6 ± 1.1
	29.5°	16	103	1	1.0 ± 1.0

earlier year at somewhat different temperatures and with different stocks (constitution of flies: Blond-Minute white/*y sn*^{3}). At 25.5° he obtained 4.6 ± 1.1 per cent of somatic crossing-over and at 29.5° 1.0 ± 1.0 per cent. Taken by themselves these values yield a difference which is not fully significant.

Discussion.—The result of these tests is that the frequency of somatic crossing-over in the X-chromosomes is low at 30° and high at both 25° and 17°. This is in contrast to most findings on the influence of temperature

on germinal crossing-over in *Drosophila melanogaster*, where the high temperature increases the frequency of crossovers as compared with the standard condition of 25°. But it should be remembered that an increase of temperature in the interval 13° to 25° can lead to a decrease of germinal crossing-over.[2]

It is interesting to compare the frequencies of mosaics known to be due to somatic crossing-over with those of mosaics due to "unstable" or "eversporting" genic loci. Gowen and Gay's data[3] with an eversporting eye color in *Drosophila melanogaster* "indicate that the incidence of eversporting varies inversely with temperature." Surrarer[4] in a striking series of observations on another unstable eye color obtained a corresponding result. Demerec[5] observed no difference in the rate of change of the unstable miniature-3 gamma gene of *Drosophila virilis* between 25° and 20°, but the value at 30° was slightly though significantly lower. Demerec interprets the lower number of mosaics at 30° as due to an observational error rather than to the effect of temperature. He points out that the wings of miniature flies reared at 30° are often wrinkled and only about half as large as those raised at 25°, so that minute mosaic spots are easily missed. In view of the results cited above it would be of interest to test the miniature-3 gamma case again and to determine if the low frequency at 30° is not due to an intrinsic cause.

The frequency of "normal" gene mutations varies directly with temperature. It is perhaps significant that the frequency of changes in "unstable" loci agrees with that of somatic crossing-over in that it varies inversely with temperature.

Summary.—The rate of somatic crossing-over between the X-chromosomes of *Drosophila melanogaster* was found not to be different at 17° and 25°, but lower at 30°. The results are compared with data from other authors on the influence of temperature on "unstable" genic loci.

[1] Stern, C., *Genetics*, **21** (1936). In press.

[2] Plough, H. H., *Jour. Exp. Zoöl.*, **24**, 147–209 (1917).

[3] Gowen, J. W., and Gay, E. H., *Science*, **77**, 312 (1933); *Genetics*, **19**, 189–208 (1934).

[4] Surrarer, T. C., *Genetics*, **20**, 357–362 (1935).

[5] Demerec, M., *Proc. Nat. Acad. Sci.*, **18**, 430–434 (1932).

STUDIES IN REPTILIAN COLOR CHANGES. I. A PRELIMINARY REPORT

BY L. H. KLEINHOLZ

BIOLOGICAL LABORATORIES, HARVARD UNIVERSITY

Communicated May 21, 1936

Although the phenomena of reptilian metachrosis have provoked the curiosity of man since the time of Aristotle, comparatively few experimental studies have been made of this subject with a view toward analyzing the physiological processes by which pigmentary activity in this group is coordinated. Most of the literature in this field consists of investigations carried on with the African chameleons (Brücke, 1852; Hogben and Mirvish, 1928; Zoond, 1934), although the publications of Parker and his associates on some of the lizards native to North America have contributed materially to our knowledge of the pigmentary activity of these forms. With the exception of the work of Carlton (1903) and of Redfield (1918), however, no attempt has been made to formulate a comprehensive physiological basis for chromatic responses in the American reptiles. Sand, in his recent publications (Zoond, 1934; Sand, 1935) takes issue with Redfield's explanation of melanophore control in *Phrynosoma* and offers an explanation of color changes based on experiments with the chameleon.

This paper is a preliminary report on the color changes of *Anolis carolinensis*, the so-called Florida chameleon. Although more detailed and extended reports are in preparation for publication in the near future, this note is offered at the present time to revise several concepts extant in the literature and to demonstrate that the mechanism for reptilian metachrosis is not uniform for the class but that it may vary with the genus of lizard used for observation.

From the studies of Carlton (1903) and of von Geldern (1921) the melanophores in *Anolis* seem to be the only cells actively concerned in the color changes of this lizard. At ordinary room temperatures (20–30°C.) specimens of Anolis are brown in the light and green in darkness. The lizard possesses a marked ability to modify its color in adaptation to the background, for if dark animals are placed upon an illuminated white background they gradually turn green, and, conversely, a green specimen when transferred to an illuminated black background rapidly becomes dark. When an otherwise normal animal is blinded, the color play in response to background disappears but the typical color reactions to light and to darkness persist apparently unimpaired. It is therefore evident that some photoreceptive element or elements in addition to the eyes are involved in the system responsible for color change.

It is well known that among the amphibians the secretion of the inter-

mediate lobe of the hypophysis is the dominant factor in pigmentary activity. When specimens of *Anolis* are hypophysectomized the ability to adapt to background changes is lost and the animals remain permanently in the green condition. The brown color may be temporarily elicited by the injection of appropriate extracts of pituitary tissue from lizards, from frogs and from fishes (Kleinholz, 1935); commercial preparations of mammalian posterior lobe (Parke, Davis & Co. obstetrical "pituitrin") are also effective in inducing darkening.

Carlton believed that the dark coloration was effected by the action of sympathetic nerve centers from which fibres connected directly to the melanophores. May (1924) studied the color changes of grafted skin in auto- and homoiotransplants and agreed with Carlton that the chromatic responses were due to the transmission of stimuli over nerve fibres supplying the melanophores. He came to this conclusion from observations that areas of skin which had been denervated (in the operation of transplantation) lost their capacity for color changes and remained green for a period of two to three weeks. At the end of this time the graft changed color when the host did. This result was attributed to the regeneration of nerve fibres into the hitherto denervated region. Neither May nor von Geldern were successful in their attempts to demonstrate the presence of innervated melanophores in histological preparations of the skin.

By making skin transplants and then testing for color changes on black and on white backgrounds it was found that grafts showed variation in their melanophore activity. In several cases of autotransplantation the transplants evidenced metachrosis on the second or third day after operation; in other cases the denervated areas did not change color until seven or eight days after this procedure. At all times, however, the transplanted area was slower than the host skin in rate of color change. These results are probably due to variations in vascularization of the grafts, since in many instances the transplanted skin shrank and the receptive area became so enlarged that good contact between the transplant and host tissue was not secured.

Homoiotransplants were made from hypophysectomized specimens to normal hosts, and the size of the transplant was trimmed to fit the receptive area of the host. Within twelve hours the transplant showed color changes upon appropriate backgrounds and lagged only slightly behind the reactions of the host skin.

Faradic stimulation (one electrode in the mouth and the second in the cloaca) of normal specimens in light and in darkness causes darkening of the skin. This darkening was shown to be due to stimulation of the pituitary, for on stimulating hypophysectomized lizards in the same manner the diffuse darkening of the skin did not occur. There appeared instead a darkening of the post-orbital regions and a scattered mottling of

dark spots over the otherwise green dorso-lateral aspect of the body and hind legs. In some animals an area of skin over the scapula behaves similarly to the post-orbital regions. The darkening of the post-orbital and scapular patches and the mottled condition also appear when the animal is excited by mechanical manipulation, such as prodding the specimen's head with a glass rod. Von Geldern has reported similar behavior of the skin of males preparing for combat.

Using the post-orbital patch as a convenient region for experimental study, it was found that darkening of this region is not dependent upon nerves, since denervation by excision and subsequent replacement of such an area has no effect on its darkening after electrical stimulation. Injection of solutions of adrenalin hydrochloride, in various concentrations, into normal dark animals causes generalized pallor with the appearance of the dark patches and "mottling"; injection of similar solutions into hypophysectomized specimens evokes "mottling" and darkening of the localized patches. Experiments are in progress to determine the effect of adrenalectomy upon the appearance of these localized colorations in response to electrical stimulation.

Summary.—The generalized darkening of *Anolis* in the light and on illuminated black backgrounds is due, not to a transmission of nervous impulses directed toward innervated melanophores, as has been previously supposed, but to an endocrine secretion from the pituitary gland. In addition, it appears that darkening of certain localized regions may be attributable to the secretion of adrenalin or an adrenalin-like substance.

Brücke, E., *Denkschr. kais. Akad. Wiss. Wien*, **4**, 179–210 (1852).
Carlton, F. C., *Proc. Amer. Acad. Arts and Sci.*, **39**, 259–277 (1903).
von Geldern, C. E., *Proc. Calif. Acad. Sci.*, **10**, 77–118 (1921).
Hogben, L. T., and Mirvish, L., *Brit. Jour. Exp. Biol.*, **5**, 295–308 (1928).
Kleinholz, L. H., *Biol. Bull.*, **59**, 379–390 (1935).
May, R. M., *Brit. Jour. Exp. Biol.*, **1**, 539–555 (1924).
Sand, A., *Biol. Rev.*, **10**, 361–382 (1935).
Zoond, A., *Phil. Trans. Roy. Soc. London*, **223**, 27–55 (1934).

AN EXPANSION-LUMINOSITY RELATION FOR NOVAE

BY F. ZWICKY

NORMAN BRIDGE LABORATORY OF PHYSICS, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated May 18, 1936

A. *Introduction.*—Novae in many respects may be considered the most interesting stellar objects. The recent recognition[1] of the existence of a special class of super-novae promises to be particularly significant. The fact that super-novae during several weeks radiate as much energy as a whole galaxy of stars suggests that observation of these unique objects will furnish valuable information on fundamental problems such as the *generation of energy* in stars, the *evolution of stars* and *stellar systems*, the *origin* and *characteristics of cosmic rays*, and the *nature of atomic rays* traveling in interstellar space.

The following lines represent an attempt to correlate some of the observations on novae and super-novae.

A complete theory must account for the *causes* of the greatly accelerated generation of energy in novae as well as for the *effects* produced by these enormous outbursts of energy. We shall here be concerned only with the latter problem.

Among the immediate effects of a sudden excessive liberation of energy in a star, the most conspicuous are the increased emission of visible light and the ejection of gaseous masses at great speeds. W. S. Adams[2] has suggested in the case of common novae that the brighter a nova the greater its velocity of expansion. The existing data, however, are hardly sufficient to establish such a relation empirically. I have therefore investigated theoretically the physical properties of some simple possible models of a nova-outburst. The results obtained suggest that we concentrate our attention on the four following observable features of novae:

(1) The absolute visual magnitude M_{max} of a nova at maximum brightness.
(2) The total energy E_{vis} which is emitted in the form of visible light.
(3) The velocity of expansion v of the ejected gases.
(4) The life-time τ of the nova.

Between these four characteristic quantities *three relations* are obtained. The theory also furnishes new information about certain characteristics of novae, which are not directly observable.

In the present paper I shall describe only the general results of the theory. In addition these results will be compared with the meager available data. A more detailed discussion of the problems involved will be published elsewhere.

B. *The Relation between the Maximum Luminosity of a Nova and Its Rate of Expansion.*—We assume in our model that, at the time $t = 0$, the energy E_0 is liberated instantaneously and uniformly throughout a sphere of radius R_0. A spherical shell of gas of total mass M which surrounds the sphere R_0 begins to expand under the action of the internal pressure. After some time the expansion will have progressed to a radius $R = R_m$ at which the gaseous shell, per radial column of unit cross-section, contains an amount of matter μ gr./cm.^{2} which is insufficient to prevent the radiation from the inside of the sphere to break through the shell and be dissipated in space. In our model R_m is approximately equal to the radius which the expanding shell has reached at the stage of maximum luminosity L_{max}. With this assumption we obtain

$$L_{max} = 3\pi R_0 \mu c v^2/2 \tag{1}$$

where c is the velocity of light. The absolute magnitude M_{max} is therefore approximately

$$M_{max} = -2.5 \log L_{max} + 88.8 \tag{2}$$

or

$$M_{max} = -5 \log v + M_1 \tag{3}$$

where

$$M_1 = 88.8 - 2.5 \log 3\pi R_0 \mu c/2. \tag{4}$$

M_1 is independent of v and should *on the average* be a constant since μ is essentially independent of the type of the nova considered. The initial radius R_0 of the star involved in the nova outburst, may, of course, vary somewhat from case to case and thus introduce individual differences in M_1.

In order to compare (3) with the observed data we tentatively adjust the value of M_1 for the case of Nova Aquilae (1918) with $v = 1.7 \times 10^8$ cm./sec. and $M_{max} = -8.8$. This leads to $M_1 = 32.35$ and

$$R_0\mu = 2.69 \times 10^{11} \text{ gr./cm.} \tag{5}$$

This value of $R_0\mu$ seems reasonable enough.

In figure 1 I have drawn in the straight line M_{max} (log v) according to the theoretical relation

$$M_{max} = -5 \log v + 32.35 \tag{6}$$

A few of the most reliable observations of M_{max} and v for various novae have been marked. The velocity v for S. Andromedae (1885) is not known from observations. Since the absolute magnitude of this star rose to about $M_{max} = -15$, equation (3) indicates a velocity of the order $v = 3 \times 10^9$ cm./sec. This value is in good agreement with the velocity of expansion

obtained independently from the expansion-energy relation described in the next section.

In the case of novae for which several absorption lines were observed at the violet end of the emission bands, the velocity of expansion given by the strongest and most persistent absorption line has been adopted to characterize the nova.

For the super-nova in N. G. C. 4273 which flared up in January, 1936, the velocity of expansion v, from our relation, was predicted to be of the order of 8500 km./sec. Subsequent observations of the spectrum of this super-nova by Humason[3] and Baade showed that for the bulk of the emitted

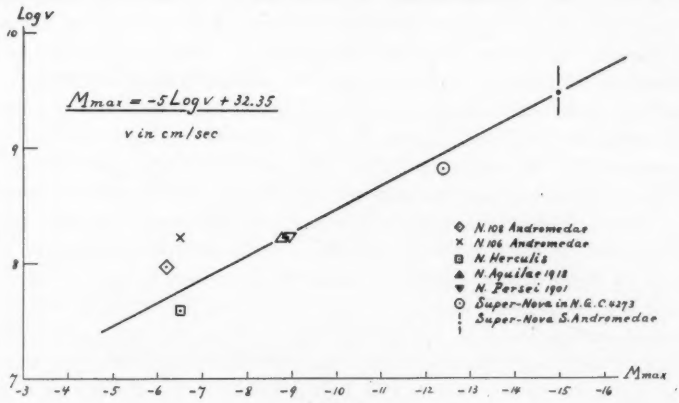

FIGURE 1

gases v was of the order of 6000 to 7000 km./sec., in good agreement with the theoretical prediction.

From the fact that our theoretical curve satisfactorily bridges the gap between common novae and super-novae we draw the important conclusion that super-novae must now definitely be thought of as giant analogues of common novae. Super-nova outbursts therefore take place in *individual* stars as Baade and I have concluded already from a different chain of reasoning.[1]

C. *A Relation between the Total Energy E Radiated from a Nova and Its Rate of Expansion v.*—The theoretical determination of the function $E(v)$ is more difficult than that of the function $M_{max}(v)$, since it requires the knowledge of the total mass M of the ejected gases. As a first approximation the following result is arrived at:

$$E = \gamma v^5 \qquad (7)$$

where γ is essentially independent of v. Also, E may be taken as the

energy E_{vis} of the visible light emitted, provided that the nova in question is not a too luminous super-nova. We therefore write

$$\log E_{vis} = 3 \log v + \alpha. \quad (8)$$

The numerical value of α can be derived only if a special model for the original star involved in the nova outburst is considered. In figure 2 the straight line representing (8) is drawn in. Adjusting the value of α for the case of Nova Aquilae 1918 we have $\alpha = 20.036$. In figure 2 the same symbols as in figure 1 are used to mark the different novae.

The visible light radiated from S. Andromedae corresponds to $E_{vis} = 1.2 \times 10^{48}$ ergs. This according to (8) indicates a velocity of expansion of

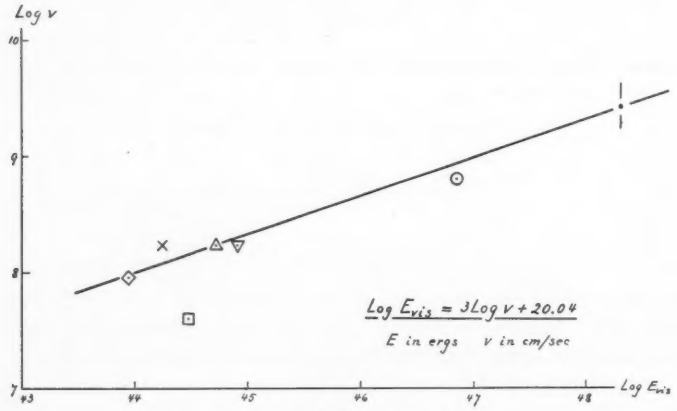

FIGURE 2

the order $v = 3 \times 10^9$ cm./sec., checking the value obtained in the previous section.

For practical purposes the use of this expansion-energy relation $E(v)$ will often be more reliable than the application of the expansion-luminosity relation. Since the light curve of many novae shows a very sharp peak, the stage of maximum luminosity can easily be missed. The total energy E_{vis}, however, can be determined with fair accuracy even from an incomplete light curve.

D. *A Possible Relation between the Life-Time of a Nova and Its Maximum Luminosity.*—The straight line of constant luminosity $M_{max} + \Delta = \text{const.}$ intersects the light curve $M(t)$ in two points whose distance in time is $t = \tau_\Delta$. We call τ_Δ the life-time of the nova as a star of magnitude brighter than $M_{max} + \Delta$. We should expect that roughly

$$E = A\tau_\Delta L_{max} \quad (8)$$

where A is approximately the same for all novae, provided that we adhere to a definite not too large value of Δ. For our simple nova-model it follows from (1) and (7) that $L_{max} \sim E_{vis}^{2/3}$ and consequently from (8)

$$L_{max} \sim \tau_\Delta^2 \text{ or} \tag{9}$$

$$M_{max} = -5 \log \tau_\Delta + \text{const.} \tag{10}$$

It is obvious that a relation of this kind would be of great value for the *determination of distances*. Assuming the validity of a relation of the type (9), approximate values for the absolute magnitude and the parallax of a nova can be obtained from a *mere inspection* of the light curve, just as in the case of Cepheid variables. Unfortunately the present data are not sufficient to formulate (9) quantitatively. It may, however, be said that the average values of τ_2 and τ_3 for super-novae are about six times as large as the averages of τ_2 and τ_3 for the four common novae N. Persei 1901, N. Geminorum 1912, N. Aquilae 1918 and N. Cygni 1920. This according to (10) indicates a difference of about four magnitudes between the brightest common novae and an average super-nova, a conclusion which is essentially correct.

In connection with the above considerations attention is called to the somewhat puzzling case of Z-Centauri, an object which flared up in the unresolved system N. G. C. 5253. The very long life time of Z-Centauri provides an additional reason for the assumption that this star was a super-nova of considerable luminosity.[4] It follows in addition that N. G. C. 5253 is an extragalactic system with a distance of the order of 10^6 light-years.

Relations of the type (9) ultimately will also enable us to decide whether or not Tycho's nova 1572, Nova Serpentarii 1604 and other new stars recorded in ancient times were super-nova.

Summarizing, I emphasize that the determination of parallaxes with the help of observations of novae can be carried out from the nearest stellar distances to the limits of the visible universe without the appearance of any "missing links." Super-novae are the *only individual stars* which can be seen at all distances in which extragalactic nebulae can be reached. Observations on novae and super-novae, therefore, promise to furnish invaluable information on every kind of cosmological problem.

E. *Suggestions for Refining the Theory.*—As stated already, a number of important effects have been neglected in our considerations. Individual differences between novae which are characterized by the liberation of equal total amounts of energy E_0 may arise for the following reasons:

(1) The variation in type of the original star. In our preliminary model we have disregarded the presence of a central star.

(2) The energy E_0 is certainly not liberated instantaneously as we have

assumed. For instance, in Nova Herculis 1934 several outbursts occurred in considerable intervals of time.

(3) The gravitational field of the original star which tends to reduce the speed of expansion.

(4) We have paid no attention to heat indices which, of course, will vary from one case to another.

(5) In general the ejected gas clouds do not constitute uniform spherical shells. Radial as well as tangential motions of the expanding gases have been observed. Complicated motions must be expected because of the action of powerful *electric propulsion* effects in addition to light pressure.

Our formulation of the three relations between E_{vis}, M_{max}, v and τ must, therefore, be regarded as a crude first approximation.

An improved theory will probably make use of the range in magnitudes $\Delta m = m_0 - m_{max}$ between the original star and the nova at maximum. Relations of the following type should then result:

$$\varphi(E, v, \Delta m) = 0 \qquad \psi(M_{max}, v, \Delta m) = 0. \qquad (11)$$

It may be hoped that the parameter Δm will take care of those individual differences between novae which arise from the differences in type of the original stars.

[1] W. Baade and F. Zwicky, These PROCEEDINGS, **20**, 254–259, and 259–263 (1934); *Phys. Review*, **46**, 76 (1934); *The Sci. Month.*, **40**, 461 (1935); see also F. I. M. Stratton, *Handbuch Astrophy.*, **7**, Erganzungsband 674 (1936).

[2] W. S. Adams, *Astrophys. Jour.*, **51**, 126 (1920).

[3] M. L. Humason, *Pubs. A. S. P.*, **48**, 110 (1936).

[4] C. Payne-Gaposchkin, *Astrophys. Jour.*, **83**, 173 (1936).